

IN THE CLAIMS:

Please amend Claims 1, 3-18, and 20-23 as follows.

---

a

1. (Currently Amended) A speech signal processing apparatus comprising:

speech segment search means for searching a speech database for speech segments that satisfy a phonetic environment;

HMM learning means for computing HMMs of phonemes on the basis of a search result of said speech segment search means;

segment recognition means for performing segment recognition of the speech segments on the basis of the HMMs of the phonemes; and

registration segment ~~determination~~ means for determining a speech segment used in the computation of the HMMs by said HMM learning means and registering the speech segment ~~to be registered~~ in a segment dictionary in accordance with a segment recognition result of said segment recognition means.

2. (Original) The apparatus according to claim 1, wherein said segment recognition means adopts diphones as units of the phonemes, categorizes speech segments into four categories CC, CV, VC, and VV (C: a consonant, V: a vowel), and performs segment recognition in each category.

3. (Currently Amended) The apparatus according to claim 1, wherein said registration segment ~~determination~~ means comprises:

pattern storage means which registers allowable speech segment patterns, and said registration segment ~~determination~~ means checks if a speech segment pattern ~~which~~ matches a speech segment that is not successfully recognized by said segment recognition means, and registers ~~that~~ the speech segment in the segment dictionary if the corresponding speech segment pattern is found.

4. (Currently Amended) ~~The apparatus according to claim 1, wherein said registration segment determination means registers~~ A speech signal processing apparatus comprising:

speech segment search means for searching a speech database for speech segments that satisfy a phonetic environment;

HMM learning means for computing HMMs of phonemes on the basis of a search result of said speech segment search means;

segment recognition means for performing segment recognition of the speech segments on the basis of the HMMs of the phonemes; and

registration segment means for determining and registering a speech segment in ~~the~~ a segment dictionary when the number of speech segments recognized by said ~~speech~~ segment recognition means is not less than a predetermined value.

5. (Currently Amended) The apparatus according to claim 4, wherein said registration segment ~~determination~~ means registers a speech segment in the segment dictionary if at least a vowel part of the speech segment is correctly recognized, even when the number of

speech segments recognized by said ~~speech~~ segment recognition means is not more than a predetermined value.

6. (Currently Amended) The apparatus according to claim 1, wherein said segment recognition means computes likelihoods of speech segments of an identical phoneme, and

said registration segment ~~determination~~ means registers, in the segment dictionary, speech segments having ~~upper~~ maximum likelihoods or having likelihoods not less than a predetermined value.

a 7. (Currently Amended) The apparatus according to claim 6, wherein said registration segment ~~determination~~ means registers, in the segment dictionary, speech segments having upper values obtained by normalizing the likelihoods by durations of the speech segments or likelihoods having the values not less than a predetermined value.

8. (Currently Amended) A speech signal processing method comprising:  
the ~~a~~ speech segment search step of searching a speech database for speech segments that satisfy a phonetic environment;  
the ~~an~~ HMM learning step of computing HMMs of phonemes on the basis of a search result of the ~~in said~~ speech segment search step;  
the ~~a~~ segment recognition step of performing segment recognition of the speech segments on the basis of the HMMs of the phonemes; and

the a registration segment ~~determination~~ step of determining a speech segment used in the computation of the HMMs in said HMM learning step and registering the speech segment ~~to be registered~~ in a segment dictionary in accordance with a segment recognition result ~~of the~~ in said segment recognition step.

9. (Currently Amended) The method according to claim 8, wherein ~~the~~ said segment recognition step adopts diphones as units of the phonemes and categorizes speech segments into four categories CC, CV, VC, and VV (C: a consonant, V: a vowel), and includes the step of performing segment recognition in each category.

a

10. (Currently Amended) The method according to claim 8, wherein ~~the~~ said registration segment ~~determination~~ step comprises:

~~the~~ a pattern storage step of registering allowable speech segment patterns, and  
~~the~~ a registration segment ~~determination~~ step includes ~~the~~ a step of checking if a speech segment pattern ~~which~~ matches a speech segment that is not successfully recognized in ~~the~~ said segment recognition step, and registering ~~that~~ the speech segment in the segment dictionary if the corresponding speech segment pattern is found.

11. (Currently Amended) ~~The method according to claim 8, wherein the registration segment determination step includes the step of registering~~ A speech signal processing method comprising:

a speech segment search step of searching a speech database for speech segments that satisfy a phonetic environment;

an HMM learning step of computing HMMs of phonemes on the basis of a search result in said speech segment search step;

a segment recognition step of performing segment recognition of the speech segments on the basis of the HMMs of the phonemes; and

a registration segment step of determining and registering a speech segment in the a segment dictionary when the number of speech segments recognized in the speech said segment recognition step is not less than a predetermined value.

a

12. (Currently Amended) The method according to claim 11, wherein ~~the~~ said registration segment ~~determination~~ step includes ~~the~~ a step of registering a speech segment in the segment dictionary if at least a vowel part of the speech segment is correctly recognized, even when the number of speech segments recognized in ~~the speech~~ said segment recognition step is not more than a predetermined value.

13. (Currently Amended) The method according to claim 8, wherein the segment recognition step includes ~~the~~ a step of computing likelihoods of speech segments of an identical phoneme, and

~~the said~~ registration segment ~~determination~~ step includes ~~the~~ a step of registering, in the segment dictionary, speech segments having ~~upper~~ maximum likelihoods or having likelihoods not less than a predetermined value.

14. (Currently Amended) The method according to claim 13, wherein ~~the~~ said registration segment ~~determination~~ step includes ~~the~~ a step of registering, in the segment dictionary, speech segments having upper values obtained by normalizing the likelihoods by durations of the speech segments or likelihoods having the values not less than a predetermined value.

15. (Currently Amended) A computer readable storage medium storing a program for implementing a the method ~~cited~~ recited in claim 8.

a  
16. (Currently Amended) A speech signal processing apparatus comprising:  
a segment dictionary in which speech segments are registered by a the method ~~cited~~ recited in claim 8;

language analysis means for performing language analysis of input text data;  
prosody generation means for generating prosody on the basis of an analysis result of said language analysis means;

speech segment selection means for ~~search~~ searching said segment dictionary on the basis of the prosody generated by said prosody generation means to select corresponding speech segments;

speech segment modification/concatenation means for modifying and concatenating the speech segments selected by said speech segment selection means; and

speech reproduction means of reproducing speech on the basis of the result modified by said speech segment modification/concatenation means.

17. (Currently Amended) A speech signal processing apparatus comprising:  
HMM learning means for ~~learning~~ learning HMMs corresponding to phonemes using a plurality of speech segments that satisfy a predetermined phonetic environment; and  
registration segment ~~determination~~ means for selecting a speech segment ~~to be registered~~ used in the learning of the HMMs by said HMM learning means and registering the speech segment in a segment dictionary used in speech synthesis on the basis of the HMMs corresponding to the phonemes.

a 18. (Currently Amended) The apparatus according to claim 17, wherein said registration segment ~~determination~~ means obtains a maximum likelihood HMM which has a maximum likelihood with one of the plurality of speech segments from the HMMs corresponding to the phonemes, checks if the one speech segment is a speech segment used in the learning of the maximum likelihood HMM, and selects the one speech segment when the one speech segment is a speech segment used in the learning of the maximum likelihood HMM.

19. (Original) The apparatus according to claim 17, further comprising speech synthesis means for producing synthetic speech using the segment dictionary.

20. (Currently Amended) A speech signal processing method comprising:  
~~the~~ an HMM learning step of ~~learning~~ learning HMMs corresponding to phonemes using a plurality of speech segments that satisfy a predetermined phonetic environment; and

the a registration segment determination step of selecting a speech segment to ~~be registered~~ used in the learning of the HMMs in said HMM learning step and registering the speech segment in a segment dictionary used in speech synthesis on the basis of the HMMs corresponding to the phonemes.

a 21. (Currently Amended) The method according to claim 20, wherein ~~the~~ said registration segment ~~determination~~ step includes the step of obtaining a maximum likelihood HMM which has a maximum likelihood with one of the plurality of speech segments from the HMMs corresponding to the phonemes, checking if the one speech segment is a speech segment used in the learning of the maximum likelihood HMM, and selecting the one speech segment when the one speech segment is a speech segment used in the learning of the maximum likelihood HMM.

22. (Currently Amended) The method according to claim 20, further comprising ~~the~~ a speech synthesis step of producing synthetic speech using the segment dictionary.

23. (Currently Amended) A computer readable program storing a program for implementing a the method ~~cited~~ recited in claim 20.